Face Recognition: a Summary of 1995 - 1997

Thomas Fromherz
International Computer Science Institute
1947 Center St, Suite 600
Berkeley, CA 94704
Phone: +1 510 642 4274 ext. 115
Email: fromherz@icsi.berkeley.edu

Keywords: Face recognition, identification, authentication, hybrid recognition, classifiers

Abstract

The development of face recognition over the past years allows an organization into three types of recognition algorithms, namely frontal, profile, and view-tolerant recognition, depending on the kind of imagery and the according recognition algorithms. While frontal recognition certainly is the classical approach, view-tolerant algorithms usually perform recognition in a more sophisticated fashion by taking into consideration some of the underlying physics, geometry, and statistics. Profile schemes as stand-alone systems have a rather marginal significance for identification. However, they are very practical either for fast coarse pre-searches of large face databases to reduce the computational load for a subsequent sophisticated algorithm, or as part of a hybrid recognition scheme.

Such hybrid approaches have a special status among face recognition systems as they combine different recognition approaches in an either serial or parallel order to overcome the shortcomings of the individual components.

1 Introduction

The following report is a summary of a longer and more comprehensive survey written in 1997 [Fromherz et al. 97], reporting on research results mostly published in the years 1995 until mid 1997, thus complementing the earlier surveys by [Samal & Igenar 92] on nonconnectionist approaches, by [Valentin et al. 94] on connectionist schemes, and the abundant survey by [Chellappa et al. 95] on 20 years of face recognition.

Modern face recognition has reached an identification rate of greater than 90% for larger databases with well-controlled pose and illumination conditions. While this is a high rate for face recognition, it is by no means comparable to using keys or batches, and not at all to a human concierge. Still, face recognition as an identification or authentication means could be successfully employed in many such tasks like entrance control, computer access control, or in the prominent field of criminal investigation. Latest news about commercial systems such as Visionics "FaceIt" even promise to bring face recognition to an extensive use in the American Department of Motor Vehicles or at ATM machines.

The development of face recognition over the past years allows an organization into three types of recognition algorithms, namely *frontal*, *profile*, and *view-tolerant* recognition, depending on both the kind of imagery (facial views) available, and on the according recognition algorithms. While frontal recognition certainly is the classical approach to tackle the problem at hand, view-tolerant algorithms usually treat it in a more sophisticated fashion by taking into consideration some of the underlying physics, geometry, and statistics.

In addition to this organization, a special status is taken by the rather novel idea of hybrid recognition algorithms, which combine different recognition approaches in either a serial or parallel order. In the parallel case submitting all single recognition rates to one of a number of possible classification procedures assesses the final results.

The following chapter gives an overview of existing frontal, profile and view-tolerant recognition approaches. The special class of hybrid recognition systems is treated in chapter 3. Some conclusions are given in chapter 4.

2 Modern Face Recognition

In this survey report, a rather "high-level" overview of face recognition is proposed by a classification into frontal, profile, and view-tolerant face recognition. It is interesting to note that the majority of direct contributions on face recognition in the two major conferences on face and gesture recognition covered in this report, 1995 in Zurich, and 1996 in Killington, chose a modern view-tolerant rather than a classical frontal approach.

In their paper about benchmark studies in face recognition, [Gutta et al. 95] independently came to the conclusion that holistic connectionist methods outperform discrete feature and correlation methods. The former are characteristic for such approaches as principal component analysis (PCA), eigenface systems or radial basis functions (RBF) and are common for view-tolerant systems while the latter use standard pattern recognition to match extracted facial features amongst faces commonly used in frontal algorithms.

The same paper also describes in detail the DARPA/ARL FERET database of facial images used in some of the research work described below. Consisting of about 1100 image sets of two frontals, and pairs of 1/4, of 1/2, of 3/4, and of full profiles, totaling about 8500 images, this image collection aims at offering a representative database for training, testing and evaluating alternative face recognition schemes.

2.1 Frontal Recognition

'Frontal recognition' denotes those approaches which in a preprocessing step find and extract facial features in head-on 2D face images which then are matched against the according features of a face database. Following facial heuristics, preprocessing steps usually are biologically motivated and serve the purpose of data reduction, removal of redundancies, and speed-up of parameter searches. Such a preprocessing step can be high-level in that it finds a list of facial features which then is searched for corresponding elements such as edge elements, lines, or circles. It can also be low-level in that filters extract and mark certain features keeping them decoded in the imagery. The subsequent matching step has the goal of finding unknown parameters such as positions and distances of facial features, and of subsequently assessing match values, which finally are used to assess an overall recognition rate.

The following methods are all specialties in the realm of frontal recognition. [Daugman 1995], for instance, describes a system where the rejection of an independence hypothesis can be used for reliable recognition. The human iris, providing the required complexity and uniqueness for this method, presents itself as a unique identifying facial feature. A real-time algorithm is presented that demodulates local phase variation in the texture using 2D Gabor wavelets, the results of which are converted into an 'Iris Code' about 173 bits long. The author claims a very high-confidence recognition by exhaustive searches at the rate of about 10,000 persons per second.

In contrast to this very biological approach, [Takács & Wechsler 96] promote a more classical perspective presenting a procedure for the representation and detection of multiscale objects based on radially non-uniform sampling lattices. The feature detection and classification involves feature encoding using visual filter banks (receptive fields) defined through self-organizing feature maps. The combination of visual filters with radially non-uniform sampling allows illumination insensitive calculation of feature codes that subsequently are collected into compact face codes for later identification.

The next contribution by [Wilder et al. 96] attempts to lessen the pose and illumination sensitivity of face recognition by testing the feasibility of infrared (IR) against visible imagery. Although the three recognition algorithms (transform coding of gray scale projection, eigenfaces, and matching pursuit filters, called in to ensure fair comparison of visible vs. IR imagery, usually would be assigned to the view-tolerant group, the comparison is based on frontal imagery recorded under fairly controlled pose, illumination, and temperature conditions. The acquisition of the imagery allowed small variations in pose, temperature, and geometry (talking and silent faces) but these were not truly corrected or normalized in any way, which is why this approach appears in this frontal recognition section.

The overall message of the experiments is that both visible and IR imagery perform similarly across algorithms, with a slight advantage of the first algorithm on IR images, and the latter two on visible images. Furthermore, recognition tests based on several facial features plus the face, without the face, and the face alone show that the face contributes to recognition in the visible case, but decreases recognition in the IR case. The results for fusing visible and IR recognition are given in chapter 3.

2.2 Profile Recognition

Facial profiles, while inherently containing less discrimination power than frontal images, are relatively easy to analyze, and, therefore, allow for fast algorithms, and show a sufficiently high number of details to support face recognition. Such a fast profile procedure can be employed as an initial search of a large face database to index candidates who then enter the more accurate but computationally expensive frontal recognition. On the other hand, profile recognition provides a simple means to support other approaches in a hybrid face recognition procedure.

Pure profile algorithms are rare. [Yu et al. 95], for instance, recently published a profile approach as a preliminary step towards a later integration into a hybrid recognition system. The procedure begins with a rule-based extraction of fiducial marks on the facial profile, which then are transformed into a one-dimensional canonical profile representation. As a preparation for the matching step, a variation model is calculated from a number of training profiles. The actual matching of a test profile with the variation models of all subjects happens by calculating the distance between the test profile points and the corresponding extremal points of the variation models.

A profile approach by [Gordon 95] in an overall valuation would belong to both the group of view-tolerant recognition and of hybrid procedures. In contrast to [Yu et al. 95], Gordon's approach extracts facial features merely to perform normalization and define template regions, which later are used for combined recognition of frontal and profile regions in a classical template matching process. The extraction techniques involve tangency constraints and heuristic knowledge about head structure.

Additionally, facial features of both frontal and profile imagery, can be used to precompute search indexes consisting of distances and angles of just these points in an attempt to speed up overall recognition.

2.3 View-Tolerant Recognition

View-tolerant algorithms describe approaches that employ various techniques to correct for perspective or pose-based effects due to illumination and to the three-dimensional nature of the head. Such techniques are mainly needed in situations where uncooperative or unaware subjects should be authenticated or identified with no view consistency given.

In a rough validation, these approaches can be organized into algorithms that construct some sort of 3D head model, into algorithms that employ Hidden Markov Model (HMM) techniques to tackle real-time recognition, or into algorithms that use some kind of deformable models and morphs to model varying shape.

The KLT usually is employed when beards, baldness, or non-uniform backgrounds are present in the imagery, since it treats the entire image, face and background, as a pattern. Computationally complex eigenface concepts are used if global and local recognition aspects shall be addressed, but only have been proved useful on a relatively small amount of images. In contrast, wavelets, mostly Gabor wavelets, extract facial features including an entire neighborhood similar to the receptive fields of the human visual system, thus representing the face as a collection of feature points.

2.3.1 Shape

All shape-based approaches need two or more images of the head from different viewing directions to produce $2^1/_2$ D or 3D models of the head. The system by [Gordon 95], mentioned in the previous section, works with one frontal and one profile image. After identifying the head bounds in the frontal view, eye candidates are extracted using general eye templates, valley based pupil detection, and then structural knowledge the human head. A similar approach is used in the profile case by first extracting the profile line and then estimating the nose and chin tip using tangency constraints and *a priori* knowledge of head structure.

After a subsequent normalization, the feature points of both frontal and profile images identify five regions, namely right and left eye, nose, mouth, and profile. These feature points enter the matching process, which consists of a cross correlation procedure with corrections for general lighting effects. Depending on the reliability and information content of the features the various regions are weighted differently, leading to a scoring of the individual matching results. It is important to notice that the profile enters the matching process by adding a limited region of the profile to the pattern matching process, and not as in [Yu et al. 95] by the geometry of the extracted feature points.

Although this method takes advantage of the shape information given in the frontal and profile images, it does not yet operate with an actual model of the head. In contrast, [Bichsel 95] presents a modular framework for shape from multiple views and varying illumination, which, in an iterative process, estimates a $2^1/_2$ D head model represented by a depth map and a corresponding texture map.

From the resulting depth and texture map any additional view of the head can be calculated within limitations. The probability estimation module, a part of the model estimation algorithm, is able to compute the probability of a specific set of shape and texture parameters, and hence could be used as a recognition module. However, since the paper mainly deals with estimating head models no recognition experiments are given.

A similar approach to shape-enhanced face recognition is given by [Fromherz 96] who proposes a shape from multiple views and visual cues procedure embedded in a framework that allows consistent integration of additional visual cues. Based on an image sequence of the head, a combination of shape from contours and shape from local luminance distribution reconstructs a head model and represents it as a set of depth maps.

The actual recognition happens by adjusting the orientation of a particular depth map together with its corresponding texture map in an iterative process until a view is reached that, in a simple template matching, shows the best fit to the image to be recognized. Alternatively, the depth and texture map could be entered in the probability module described above ([Bichsel 95]) to receive a probability value for the according face model.

In order to use even more accurate head models, [O'Toole et al. 95] employs range data recorded with a Cyberware rotational scanner to calculate an average head model represented by range and texture data. Assuming that the viewing direction of a recorded image is known, the model is used to compute new views of the head. The paper sets out to show that overlapping visible regions of heads can support accurate recognition even with large viewpoint changes.

While the approaches mentioned so far work with images recorded under fairly controlled conditions, [Clergue et al. 95] work with multimedia documents that contain uncontrolled video segments from which an analysis of the body motion and the identity of persons in motion should be accomplished. This initial paper shows the basics of how to achieve identification based on projective invariants and on model-based estimations of head pose. In contrast to the methods above, which tried to find an accurate head model of a subject, this approach works with a generic 3D-mesh model of a human body.

2.3.2 Hidden Markov Modeling (HMM)

While Hidden Markov Models have been used in speech recognition for more than a decade, and were also promoted for gesture recognition in recent years, only little work has been done on applying HMM to face recognition ([Samaria & Young 94], [Nyffenegger 95], [Achermann & Bunke 96]).

HMM generally works on sequences of coherent 1D signals (feature vectors), while an image usually is represented by a simple 2D matrix. This dilemma can be solved by applying a sliding window to the image covering the entire width of the image, which is moved from the top to the bottom of the image. The brightness values of the windows are passed to the HMM process as 1D-feature vectors. Successive windows overlap to avoid cutting of significant facial features and to bring the missing context information into the sequence of feature vectors. The human face can be divided in horizontal regions like forehead, eyes, nose, etc. that are recognizable even when observed in isolation. Thus the face is modeled as a linear left-right HMM model of five states, namely forehead, eyes, nose, mouth, and chin.

The complete recognition system incorporates a preprocessing step followed by four processing steps: (1) The feature vectors of a test set of images is used to build a code book; (2) Based on this code book all feature vectors are quantized; (3) The HMM model is trained for every person in the database; (4) Recognition: A test image, not used for training, passes the preprocessing step and step 1 and 2, before the probability of producing this image is computed for every person, *i.e.* every model, in the database. This procedure produces a ranking of the potential individuals in descending order of the probability of each model.

2.3.3 Deformable Models

In contrast to the 'physical' approach of shape-based face recognition, the idea of deformable models is to encode regions around significant facial features in a mathematical description and let a similarity function decide about the relations among the descriptions of different faces. Since it is impossible to include all possible parameters of facial imagery in a face recognition model, data driven methods are employed to design key components of the model.

One such method, used by various authors to automatically extract facial features, employs a wavelet decomposition of the images. [Phillips & Vardi 95], for instance, present an adapted wavelet decomposition as key element of their matching pursuit filters to find the subtle differences between faces. The ability of wavelets to encode local information minimizes the effect of variations of background and facial expressions.

In this face recognition system two sets of matching pursuit filters play the key role of locating facial features and of identifying unknown faces. The first set of filters determines the face position in the image and reports the locations of nose and eyes, which are split up into five facial features represented by a 5-component vector. The identification happens by computing a similarity score for each feature of the probe (unknown face) and the corresponding feature of one candidate of the gallery (known faces), followed by the computation of a total score for this candidate. Repeating this procedure for all images in the gallery finally leads to the identification of the probe.

[Phillips & Vardi 95] also present another application of data-driven algorithms, illumination normalization. The presented algorithm corrects for both intra-face variations where a face is not uniformly illuminated, and inter-face variations where the illumination of one face is normalized to a second face of standard illumination so that the two faces can be compared. This happens by transforming the histogram of one image into the histogram of the other image employing a general form of histogram equalization.

Data-driven algorithms can be found in the majority of the deformable model approaches. [Konen & Schulze 95], [Wiskott et al. 95], and [Maurer & Malsburg 95], for instance, employ an elastic graph method that stores the faces as grids (graphs) with the characteristic facial features attached to the nodes of the graphs. The features are obtained by convolutions of the faces with Gabor wavelets computed at the node locations. Through the Gabor wavelets not only local features but an entire neighborhood of the features are integrated.

Once the features have been extracted the creation of a graph happens by matching a stored graph to the actual image by an optimization procedure where position, size, and inner structure of the graph are elastically varied to maximize the similarity of graph and image features. The resulting labeled graph is then compared with a number of precomputed reference graphs, the general face knowledge, and only if certain significance conditions are fulfilled, the match is accepted. The elastic graph method is robust with respect to variations in pose, size, and facial expression of the head and can deal with different lighting conditions.

The real world access control system 'ZN-face' reported by [Konen & Schulze 95] employs user-triggered semi-automatic image acquisition, allows fast identification in about 3.5 seconds, and can handle large databases of more than 1000 subjects due to verifying instead of recognizing a subject's identity.

[Wiskott et al. 95] modified the elastic graph approach so that the nodes in the different graphs always refer to the same local facial region. This allows one to enhance face recognition by additionally detecting certain features like gender, beards, or glasses in the image of the recognized person. This is accomplished by checking whether the best fitting graph in the general face knowledge is mostly male or female and did or did not show evidence for beard or glasses.

[Maurer & Malsburg 95] tackle the problem of single-view based recognition across varying view points by introducing a geometric transformation which acts on the encoded feature points (so-called jets) in the graph. Thus, if the surface normals are given at the graph nodes and if the angle between two views is known, then the graphs can be transformed by simply multiplying every jet with its geometric transformation matrix. Assuming a fixed set of normals, the required pose and normals are found by optimizing recognition on a training set of frontal and half profile views.

While the above methods are merely robust against variations in lighting, more sophisticated approaches decompose geometry and lighting to achieve better results under varying contrast. The contribution of [Craw et al. 95], for instance, decomposes a number of faces into configuration vectors (shape vectors) and shape-free faces (texture vectors). From an ensemble of faces, manually coded into feature vectors containing the 34 most significant facial landmarks, shape variations ordered by importance are extracted into a new independent eigenface basis. It is now easy to project every face onto this basis and compute the components of the face along every eigenface.

In the shape-free approach, each face of the ensemble is texture mapped to a standard shape, the average shape of the set of ensemble images. A principal component analysis then decomposes the texture mapped shape-free faces into eigenvectors describing the lighting variations. Again the components of every face along these eigenvectors are computed. Finally, recognition of a coded test image is achieved by matching the component vector of the test image against the ensemble components, for instance, using an Euclidean metric.

Another approach to recognition from a single view is given by [Lando & Edelman 95] who choose a representation of face images similar to [Konen & Schulze 95] and [Takács & Wechsler 96]. Following biological visual systems where the sensitive regions of a neuronal unit is known as *receptive field*, the employed model represents face images by vectors of activities of graded overlapping receptive fields (RF).

Same-class objects like faces induce similar patterns (clustering) in RF space across varying viewing conditions (pose and illumination) and facial expression. [Lando & Edelman 95] propose a face recognition scheme consisting of three modules. The first employs radial basis function (RBF) classifiers, which accepts image representations in the high-frequency RF space thus detecting the viewing conditions. In a class-specific transformation that corresponds to the detected viewing conditions, the second module transforms representations in the low-frequency RF space into a prototype representation predicted for the input image. In the final module, face identification happens by RBF classifiers, which in the low-frequency space compares this predicted representation with those of known faces. Since both classifiers need training on two databases, which can be updated with the resulting class-specific transformation, the system is capable of improving with experience.

A special position among deformable model approaches is taken by morphing techniques that employ various methods to find suitable parameters (so-called morph fields) to represent the linear transition of one image into another. In this context, the recent algorithm by [Bichsel 96] generates optimum image morph fields based on distortion of geometry and illumination by maximizing the probability of the morph fields in a Bayesian framework. The method needs no training and is derived based on invariances and basic transformation groups.

The author shows that, assuming rigid objects with Lambertian surfaces without self-shadowing and ignoring occlusions, all possible morphs among images of a single person are confined in a 5-dimensional subspace of the overall morph space. Although these constraints are not always applicable to faces one can assume that the five most significant eigenvectors of the morph measurement matrix, calculated from at least three images, govern the major inter-personal variations.

Recognition now happens by computing the morphs from the unknown image to the reference images of all possible candidates, by then projecting the resulting morphs into the five-dimensional subspace of each person, and by finally using the length of the residual vector as match value.

Tests show excellent morphing results for morphs between two persons of different view and illumination. Although no recognition results are given, preliminary recognition tests show that morph fields are well suited for recognition under varying view and illumination.

3 Hybrid Systems

An area that has received significant attention in recent years is classifier combination and sensor fusion. Realized in so-called hybrid systems, sensor fusion is already commonly believed to be of advantage in shape reconstruction from video images ([Fromherz & Bichsel 95]), in optical character recognition, and in the combination of face and speech recognition for person identification. In a recent study of automatic face recognition systems, [Gutta et al. 95] finally came to the conclusion that the future in face recognition lies in the research of hybrid recognition systems.

Different recognition approaches succeed and fail at widely different viewing and illumination conditions, which usually cannot be accommodated by day-to-day recognition systems. Due to this dilemma, it seems obvious to run various individual recognition classifiers on a problem leading to an individual ranking of the results of every process, and to design a classification scheme to assess an overall recognition result. In contrast to this parallel approach of equivalent face classifiers also a serial one is conceivable where the output of one classifier is input to the next one. The latter approach can even go along with hybrid learning if the classifiers require training.

As mentioned in section 2.3.1, the recognition method by [Gordon 95] already could be counted as a first attempt to build a hybrid system combining template-based frontal and profile recognition. Certainly, Gordon's profile approach only adds an additional template to the matching of the frontal templates and, therefore, does not constitute an independent recognition method. Still, the overall template matching is subject to a scoring of the five facial templates.

Another recognition project that can be counted as a hybrid recognition attempt is the work by [Wilder et al. 96] on visible and IR imagery already presented in section 2.1. It has to be noticed, however, that the term 'hybrid' refers to the combination of different imagery rather than of different recognition approaches and that the combination does not involve any ranking or classification scheme among the two image types.

An example of a proper serial hybrid classification scheme involving hybrid learning was presented by [Gutta et al. 96] who propose that intelligent hybrid systems include specific levels of knowledge, namely connectionist levels which can handle different sensory input and symbolic levels which are able to fuse data from different sensory modalities and cognitive modes. The presented system for face and hand gesture recognition combines ensembles of radial basis functions (ERBF), a holistic template matching able to cluster similar images before classification, as the connectionist level with inductive decision trees (DT), an abstractive matching using discrete features, implementing the symbolic level.

[Gutta et al. 96] proposed and tested two different ERBF architectures. ERBF1 separately trains the same three RBF nodes on three different sets of images, the original images, the original images distorted by Gaussian noise, and the original images distorted by a geometric transformation (rotation). ERBF2, on the other hand, trains the same three RBF nodes on a combination of the imagery of ERBF1. The output of both versions consists of objects described by a fixed set of attributes and their discrete attribute values assigning each objects to either of two classes. The goal of the following symbolic stage is to derive rules (the DT) for classifying these objects based on a collection of training objects with known class labels.

Two different experiments were conducted, the first an identification task like 'find person X with/without glasses', and the second a verification task, for instance, for forensic verification. The first experiment is implemented in two hybrid stages: the match stage (only involving an original RBF network) finds the identity of the probe, while the second stage (DT) checks for the presence or absence of glasses. In the forensic experiment a large number of candidates are tested against a predefined image gallery on which the system has already been trained. The hybrid character in this case also involves that the training inputs for the DT method are generated from the already trained ERBF method. The results show that ERBF architectures outperform simple RBF approaches, that hybrid learning improves classification performance, and that training on a combination of original and distorted data (ERBF2) leads to improved performance against training on separate sets of training data (ERBF1).

A proper parallel hybrid recognition system integrating three face classifiers, namely the profile approach by [Yu et al. 95], an HMM algorithm similar to the one in [Samaria & Young 94], and the eigenface approach by [Turk & Pentland 91], was successfully presented by [Achermann & Bunke 96]. All three face classifiers return a ranking of the possible person in ascending order of its score. For the subsequent combination classifier, combining the results of the individual classifiers, three decision strategies, *voting*, *ranking*, and *scoring* are conceivable. With *voting*, every classifier has one vote and a decision is reached through the majority of the votes. With the second strategie, on the other hand, a ranking for the combination classifier is computed, that depends on the ranks assigned by the individual classifiers. And finally with *scoring*, information of the score function of the individual classifier is used to assess a ranking for the combination classifier. Of all three strategies, the score-based is usually favored since it provides additional information not available through voting or ranking.

Experiments show that the different recognition criteria of different face recognition approaches can be effectively combined to enhance the identification performance that can be achieved with a given image gallery. Tests also show that for the test images the combination of three face classifiers is superior to two classifiers. Future research, however, has to investigate whether this is a general behavior and how much influence comes from the quality of the images and the choice of the face classifiers.

4 Conclusions

A classification of face recognition algorithms into frontal, profile, and view-tolerant approaches was given. Naturally emerging from the imperfections of these algorithms, and already suggested by various authors, is the need to employ hybrid recognition systems that classify and combine the results of different algorithms. In this survey, several different hybrid approaches have been listed and discussed.

Most of the recognition systems available today have been demonstrated only on limited datasets, often recorded under restricted conditions. This is legitimate if commercial applications with special recording constraints are produced, or if completely new recognition algorithms are examined. In the research of face recognition, however, more often than today advantage should be taken of large public databases like the FERET database, to allow comparable results, and to advance recognition beyond the point of specialized environments.

Finally, future algorithms should add another classification level to the existing ones, namely the level of techniques that can handle nonrigid facial motions (facial expressions) and unconstrained real-time video sequences.

5 Bibliography

[Achermann & Bunke 96] Bernard Achermann, Horst Bunke, “Combination of Face Classifiers for Person Identification”, *Proceedings of the 13th IAPR International Conference on Pattern Recognition (ICPR)*, 3:416-420, Vienna, 1996.

[Bichsel 95] Martin Bichsel, “Human Face Recognition: From Views to Models - From Models to Views”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition, IWAFGR 95*, 59-64, Zurich, 1995.

[Bichsel 96] Martin Bichsel, “Automatic Interpolation and Recognition of Face Images by Morphing”, *Proceedings of the International Conference on Automatic Face and Gesture Recognition, ICAFGR 96*, 128-135, Killington, 1996.

[Chellappa et al. 95] Rama Chellappa, Charles L. Wilson, Saad Sirohey, “Human and Machine Recognition of Faces: A Survey”, *Proceedings of the IEEE*, 83(5):704-740, 1995.

[Clergue et al. 95] E. Clergue, M. Goldberg, N. Madrane, B. Merialdo, “Automatic Face and Gestual Recognition for Video Indexing”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition, IWAFGR 95*, 110-115, Zurich, 1995.

[Craw et al. 95] Ian Craw, Nicholas Costen, Takashi Kato, Graham Robertson, Shigeru Akamatsu, “Automatic Face Recognition: Combining Configuration and Texture”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition, IWAFGR 95*, 53-58, Zurich, 1995.

[Daugman 95] John Daugman, “Face Recognition by Feature Demodulation”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition, IWAFGR 95*, 350-355, Zurich, 1995.

[Fromherz & Bichsel 95] Thomas Fromherz, Martin Bichsel, “Shape from Multiple Cues: Integrating Local Brightness Information”, *Proceedings of the Fourth International Conference for Young Computer Scientists, ICYCS 95*, 855-862, Beijing, P.R. China, 1995.

[Fromherz et al. 97] Thomas Fromherz, Peter Stucki, Martin Bichsel “A Survey of Face Recognition”, *MML Technical Report, No 97.01*, Dept. of Computer Science, University of Zurich, Zurich, 1997.

[Fromherz 96] Thomas Fromherz, “Shape from Multiple Cues for 3D-Enhanced Face Recognition: A Contribution to Error Reduction by Employing Inexpensive 3D Reconstruction Methods”, *Ph.D. Thesis*, University of Zurich, Zurich, 1996.

[Gordon 95] Gaile G. Gordon, “Face Recognition from Frontal and Profile Views”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 47-52, Zurich, 1995.

[Gutta et al. 95] Srinivas Gutta, Jeffrey Huang, Dig Singh, Imran Shah, “Benchmark Studies on Face Recognition”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 227-231, Zurich, 1995.

[Gutta et al. 96] Srinivas Gutta, Jeffrey Huang, Ibrahim F. Imam, Harry Wechsler, “Face and Hand Gesture Recognition Using Hybrid Classifiers”, *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, *ICAFGR 96*, 164-169, Killington, 1996.

[Konen & Schulze 95] Wolfgang Konen, Ekkehard Schulze-Krüger, “ZN-Face: A System for Access Control Using Automated Face Recognition”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 18-23, Zurich, 1995.

[Lando & Edelman 95] Maria Lando, Shimon Edelman, “Generalization from a Single View in Face Recognition”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 80-85, Zurich, 1995.

[Maurer & Malsburg 95] Thomas Maurer, Christoph von der Malsburg, “Single-View Based Recognition of Faces Rotated in Depth”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition, IWAFGR 95*, 248-253, Zurich, 1995.

[Nyffenegger 95] C. Nyffenegger, “Gesichtserkennung mit Hidden-Markov-Modellen”, *Masters Thesis*, IAM, University Bern, 1995.

[O’Toole et al. 95] Alice J. O’Toole, Heinrich H. Bülthoff, Nikolaus F. Troje, Thomas Vetter, “Face Recognition across Large Viewpoint Changes”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 326-331, Zurich, 1995.

[Phillips & Vardi 95] P. Jonathon Phillips, Yehuda Vardi, “Data-Driven Methods in Face Recognition”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 65-69, Zurich, 1995.

[Samal & Iyengar 92] Ashok Samal, Prasana A. Iyengar, “Automatic Recognition and Analysis of Human Faces and Facial Expressions: a Survey”, *Pattern Recognition*, 25(1):65-77, 1992.

[Samaria & Young 94] F. Samaria, S. Young, “HMM-Based Architecture for Face Identification”, *Image and Vision Computing*, 12(8):537-543, 1994.

[Takács & Wechsler 96] Barnabás Takács, Harry Wechsler, “Visual Filters for Face Recognition”, *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, *ICAFGR 96*, 218-223, Killington, 1996.

[Turk & Pentland 91] Michael Turk, Alex Pentland, “Eigenfaces for Recognition”, *Journal of Cognitive Neuroscience*, 3(1):71-86, 1991.

[Valentin et al. 94] Dominique Valentin, Hervé Abdi, Alice J. O' Toole, Garrison W. Cottrell, “Connectionist Models of Face Processing: a Survey”, *Pattern Recognition*, 27:1209-1230, 1994.

[Wilder et al. 96] Joseph Wilder, P. Jonathon Phillips, Cunhong Jiang, Stephen Wiener, “Comparison of Visible and Infra-Red Imagery for Face Recognition”, *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, *ICAFGR 96*, 182-187, Killington, 1996.

[Wiskott et al. 95] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, Christoph von der Malsburg, “Face Recognition and Gender Determination”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 92-97, Zurich, 1995.

[Yu et al. 95] Keren Yu, Xiaoyi Jiang, Horst Bunke, “Face Recognition by Facial Profile Analysis”, *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, *IWAFGR 95*, 208-213, Zurich, 1995.